17 | 招聘篇：如何写一个合适的JD？

2022-03-09 DS Hunter

《反爬虫兵法演绎20讲》 课程介绍 >

讲述：DS Hunter
时长 18:05 大小 16.57M

你好，我是 DS Hunter，反爬虫专家，又见面了。

我在开篇词里说过，学完前两章的你，单兵作战的能力已经达到了极致。那么现在，我们就一起开始从单兵作战到团战的跨越。或者，你也可以理解为我们要实现从技术到管理团队的跃迁，**开启反爬虫团队作战的新模式。**

反爬虫搭建团队，第一件事就是招人。而招人的第一件事，就是编写一个合格的 JD。

如果你曾经做过管理，你可能会说：这个我熟，JD 如何如何写，能够写得精准，给应...更好的感受，寻找到匹配的人才。

你的经验是非常正确的。但是我们来看这样一个故事：

> 一个冷风嗖嗖的春天，一个朋友找到我：“哥，帮个忙，我们被竞对爬的不行了，他们技术太牛逼了，我们检测了各种框架都被绕掉了，你有办法帮我们看下他们的规律不？我把日志权限开给你。”
>
> 我说：“是你们的 XX 竞对吗？”
>
> “是的。”
>
> “稍等我十分钟。”
>
> 十分钟后，我给了他们一份文档，里面详细写了对方使用了什么技术，后端什么样子，破解用的什么，人员组成是什么样的。
>
> 朋友大惊：“你这太强了，怎么做到的，我日志里都找不出这么全的信息！”
>
> 我立刻甩给他一份 JD——他们的竞对正在招聘爬虫工程师，长长列出了各种需求，甚至根据薪资情况还能推出对方能招到什么样的人。
>
> 朋友不屑地说：“切，我还以为多厉害呢。”

你看，反爬就是这么憋屈，你的每一个操作都像一个魔术，一旦揭秘了，对方就会觉得：哦，不过如此。

然而过了不到一周：

> “哥，我们这面反爬团队也搭了半年了，至今也找不到什么人啊。你看有没有办法帮我看看？”
>
> 我又甩出一份 JD 给他。这是他们自己在网上发布招聘的 JD，基本啥都是面议。我说：“你写这么点谁知道你们要干啥啊，这有人应聘就出鬼了。”
>
> “你说得对”，朋友恍然大悟，说，“我马上去改改 JD，这谁写的 JD，太业余了。我得把我们的要求全写出来，比如我们需要……咦等等，是不是……”

> 朋友忽然好像又想起了什么。他翻出之前我发给他的竞对 JD，陷入了沉思……

看吧，现在你还觉得 JD 好写么？

现在各个厂的 JD 其实很混乱，小厂么写得详细无比，因为他希望你入职就能直接干活。大厂嘛，JD 跟写着玩似的，根本不在乎你之前做没做过，因为他有培训机制，只需要优秀的人才即可。

爬虫和反爬虫领域，也有类似的问题，只不过小厂通常没有反爬虫，因此少了一种情况。所以我们要写的 JD 有三种：大厂爬虫，大厂反爬虫，小厂爬虫。后续我们会分开讨论。我们先看下 JD 的通用规则。

JD 通用规则

JD 是候选人接触公司的第一窗口，它的好坏直接决定了应聘者对你的第一印象。一般说来，JD 的要求是：内容丰富、精准、简洁。

具体内容主要包括：

1. 岗位工作内容，职责，汇报线（可选）
2. 任职必备资格
3. 技能以及经验要求
4. 薪资结构区间

其中第二点，慢慢在淡化。由于学历歧视的问题，越来越多的公司开始在 JD 里去掉学历要求。同理，还有年龄要求，性别要求等等。当然，这些要求可能并未去除，只是变成了潜规则。慢慢的，似乎就没什么必备资格可言了。因为随便写点什么都可能引发 PR 问题，那还不如不写。

JD 的初始写作者通常是技术人员，然后 HR 做审核。有些技术人员可能觉得 HR 审核的时候没有改什么，但是实际上 HR 的修改通常比较核心。例如你有学历要求，他帮你去掉了，一定不要和他争。不要担心找来的人不符合要求，要知道，第一，你们是有面试的；第二，JD 通常分内外版本，外部版本是给候选人看的，内部版本才是用来筛选的，这是常规操作——你的反爬信息收集不还分爬虫可见版本和内部解密版本吗？

大厂 JD 潜规则

大厂的 JD 跟玩似的。

这真不是黑大厂，你去看看线上 JD 就会发现小厂写得还挺详细的，大厂的 JD 就跟随便写着玩似的，甚至有 HR 坦诚清晰地和你说：不用看 JD，那就是随便写的。很多初次面试大厂的人都对此很不习惯。

其实大厂的 HR 也很为难，JD 写太详细了，容易被人挑毛病。不如写得含含糊糊的，之后万一想拒了你，也可以说不符合要求。如果写详细了，容易被人杠：你这么多要求，我全符合，你凭啥说我不符合要求。但是少写点，解释起来就可以天马行空了。

举个例子：要求 Vue 熟练。

你 Vue 特别厉害，怎么考也考不倒，但是另一个候选人接了 offer，必须拒绝你，就很难找理由。但是如果 JD 写：要求 Web 主流框架熟练。Vue 你考不倒还有 React 吧？不行问 jQuery，总有办法的。

出于这种潜规则，你的 JD 本身也要符合这样的要求。此外，记不记得我们前面的故事里，JD 太详细，会被别人直接抓出技术栈？所以，这也是出于行业的特殊性，不得不这么干啊！

小厂 JD 潜规则

小厂的潜规则则刚好相反。

小厂自由奔放很多，不用考虑那么多 PR 相关的事情，极端的特别小的厂，逼急了甚至都敢倒闭。可是，针对人员要求，他们的要求则极为精准，恨不得连你之前用哪个框架都有限制，巴不得天下候选人都是为自己量身定制的。

此外，小厂对大厂候选人有时候迷之崇拜，有时候又坚决不用——当然，大部分情况下是因为薪资接不住。

爬虫反爬虫 JD 分类讨论

了解了大厂和小厂的潜规则之后，我们 JD 的改进方向就有了。我们按照大厂和小厂的爬虫、反爬虫细分下，分别来看。

大厂反爬虫

首先第一个，大厂的反爬虫。

大厂的优势是，只要应聘者聪明，一切皆可培训。那么大厂写 JD 这件事其实就很简单了。这个时候，他有两个选择：

1. JD 直接要求招反爬虫的人
2. 自己不写 JD，在安全和前端的求职者里面去捞人

首先第一个，自己写 JD 招人。秉承着大厂一切含糊处理的精神，这里举一个反爬虫的例子：

> 了解前端反爬虫对抗，有实际对抗经验

这里并没有说用什么办法对抗，也没有说我们希望如何对抗。是重规则引擎，还是重前端加密？实际上，都是可以的。

事实上，这类 JD 只是用于初步筛选一下应聘者的技术栈点在了哪里。从前面所有课程我们可以看到，反爬虫的天赋技能学习，主要是点前端的，但是顺便也要点一部分后端。因此两部分都是有要求的。

那么，在上面这句话的要求下，其实还需要再补充一句：

“针对数据库访问，缓存以及日志系统有自己的见解”

为什么一定要提自己的见解？

因为后续 HR 在拒绝面试人的时候，如果面试人不依不饶非要给个说法，HR 还可以在“自己的见解”上面发挥空间。当然，这只是用于兜底的，大部分候选人都很忙，同时在面好多家，没有时间和 HR 扯。不要妄自尊大觉得自己同时面了很多求职者，就觉得自己牛了。实际上，求职者也同时在面多家。这是一个多对多的关系。

最后关于薪资。大厂薪资其实相对是比较透明的，没有太多的议价空间，大家也不是很在乎。最终，大家还是以面试结果定级为准。

大厂爬虫

好，终于到了大厂爬虫这个十分尴尬的分支了。

十分讽刺的是，大厂通常不对外承认自己有爬虫团队——虽然为了避免热度，也不特意去否认。不过，他们确实也在招聘平台上放出了大量的爬虫工程师岗位。目前有的大厂已经在注意这类岗位带来的 PR 风险并尽量避免了，而有的大厂却依然不当回事，直接挂着职位。

其实，正确的做法应该是在反爬虫的 JD 里顺便招爬虫工程师。不过，那样效率太低了，大部分情况下，大厂还是在直接招聘。所以我们还是来看一下这种情况下的 JD。

举个例子。随手找到一个 JD，就不列举公司名称了。它是这么写的：

> 五年以上爬虫工作经验，精通 Java 开发
> 丰富的电商数据爬取经验
> 熟悉各类反爬虫策略，并有丰富的反反爬虫实战经验

这段 JD 直接告诉了全世界：他们在电商领域有个竞对，并且他们在爬竞对数据。

那么如何修正这个问题呢？

很简单，第二条直接删掉即可。这样即使你在招爬虫工程师，别人也不知道你要爬谁。你完全可以说：我们有个合作伙伴，因为 API 来不及开发了，不得不用爬虫来做。虽然大家都知道你在撒谎，你也知道大家都知道你在撒谎，可是你就是可以继续撒谎，能怎么样呢？

何况，电商爬虫属于通用爬虫，电商本身并没有什么与众不同的地方。这一条完全是多此一举。

除此之外，这个 JD 还能分析出什么？

是的，他们使用的爬虫框架是 Java。没错，不是 Python，不是 Node，是 Java，这个非主流的爬虫语言。

一旦锁定了语言，那么他使用了什么框架，范围就缩小了很多。你可以逐步检测特征，慢慢就能检测到他用的是什么框架了。如果不说是什么语言，全世界所有的框架一个一个试，你要试很久了。

那么如何修正这个问题呢？

有两个办法，一个依然是删掉，另一个是，有意指定一个错误的语言来误导对方。

所以呢，你可能会问我，会不会他们实际上使用的并不是 Java 呢？这个问题提的非常赞，虽然我也不能保证什么，不过，只能说，试试检测 Java，并不亏哦！

我们再来看一个正面例子：

> 负责爬虫系统的核心技术研发，支持海量数据采集分析，实现业务赋能。
> 负责互联网资源特征挖掘，包括页面特征，价值，变化趋势等等
> 设计各种策略与算法，提升爬虫系统资源收录与更新效果。

你看，这个 JD 和上面的 JD 比起来，就更“大厂”了一些，不但啥细节也没透露，还说了一大堆没用的废话。换言之，不说人话，这正是大厂在 PR 压力下的生存之道。

小厂爬虫

小厂的爬虫就嚣张跋扈了很多。

我们刚刚提到 JD 提爬虫会引发风险。注意，这里的风险大部分不是法务风险，仅仅是 PR 风险而已。要知道，走法务流程，证据是十分重要的。你不能说我发布了一个 JD 说要招爬虫工程师，那就证明我已经有爬虫团队了对吧？更不能说我已经在爬竞对了。因此，小厂的 JD 你可以尽可能精确要求候选人需要什么技能。

还是举个例子：

> 一年以上爬虫相关经验，Python 优先

> *（实际上他们就是在用*Python）*
>
> 熟悉各种抓取技术，熟练使用正则表达式从各网页中提取信息
> *（他们真的在用正则！因为正则真的很快！）*
>
> 了解 JS 逆向以及 App 逆向
> *（说明他们同时抓竞对两个版本，很可能还存在跨平台比价）*
>
> 对用户增长理念认同者优先。

这个 JD 我们可以得出结论：他们在用 Python 抓取竞对，同时抓取 JS 和 APP，而且很可能存在跨平台对比。同时，因为使用的是正则，要么他们爬虫水平特别棒，要么竞对反爬特别不给力。

但是需要修改吗？不需要！

小厂怕啥啊，逼急了敢倒闭的主，你拿个 JD 去起诉人家，法官都要笑话你啊。

这样精准的描述可以快速拿到合适的人选。虽然反爬虫方搜一下你的 JD 就知道你的玩法了，但是一般说来，他也没什么办法啊，不然也不会沦落到用正则就能直接提取出网页信息的地步。如果爬虫能被逼到起浏览器抓取，那怎么着 JD 得要求 xPath 能力吧？或者其余选择器。

不过如果想再优化一下，也是可以的。你可以多加点信息，然后面试的时候再筛一下人。例如，你可以说熟悉 Python，Java，XXX……等等语言的优先。即使这些语言你不用，大不了简历筛选不通过嘛。就算简历筛选过了，还有面试呢，一面就能拦住，不要虚。

加餐：小厂反爬虫

这个加餐加得我非常纠结。

因为原则上来说，小厂很少做反爬虫。能玩得起反爬虫的，基本上都成为大厂了，至少是垂直领域的老大。所以，小厂反爬虫……真的在这个世界上存在吗？

但是很多小厂找我寻求过技术支持。因此可见，他们虽然玩不起，但是还是存在这个需求的。

但是小厂的反爬虫真的出不起成本。因此我的建议是，不如干脆让爬虫 leader 兼职反爬虫，带一个人顺手做做就好了。反爬本来就是一个烧钱的项目，ROI 又极低，小厂，尤其是创业公司，本来现金就非常宝贵，干嘛非要砸这个钱呢？

因此，需要在爬虫的 JD 里面加一句：“能同时实现反爬虫需求”。这就足够了。

自驱

自驱这个词已经被一些大厂玩的快成为贬义词了。但是，这个词本身的含义是很正面的。假设你是一个爬虫或者反爬虫的 Leader，在你主动出击寻找对应的人才的时候，完全可以因人设岗。这个情况下，你就不需要写 JD 了。因此，你也就不用为这个问题苦恼了。

嗯，是的，我曾经去面试过一个反爬虫的岗位。他们的 JD，写的是：前端开发工程师……

小结

这一节，我们主要探讨了大厂与小厂，爬虫与反爬虫，所有情况下的 JD 写法。

由于大厂对 PR 风险考虑更高，小厂对人员精准、降低用人不匹配问题的要求更高，因此大家各有各的写法。大厂更偏向于通用描述，小厂更偏向于详细要求每一个细节。

总体来说，反爬虫 JD 可以写得细致，而爬虫 JD 则要考虑法务风险，尽可能不要提“破解”，更要尽量做到不暴露自己的技术栈与框架，这是几个基本原则。

思考题

好了，又到了愉快的思考题时间。还是老规矩，三选一。

1. 一个公司，发了一个 JD，要求应聘者同时做过爬虫和反爬虫，你认为他们实际上想招哪一个？
2. 你和 HR 说要招一个爬虫工程师，HR 直接去网上搜了个 JD。这样最坏可能引发什么问题？

3. 如果真的用 JD 做证据起诉对方爬虫团队，居然还赢了，那么你觉得最大的可能性是爬虫团队犯了什么错误？

期待你在评论区的分享，我会及时回复你。反爬无定式，我们一起探索。

极客时间

招兵买马举大义，卧薪尝胆立功名。

反爬虫兵法演绎 20 讲

DS Hunter / 反爬虫专家，前某公司研发总监

识别二维码
免费试读

分享给需要的人，Ta订阅超级会员，你将得 50 元

Ta单独购买本课程，你将得 20 元

生成海报开分享

赞 1 提建议

© 版权归极客邦科技所有，未经许可不得传播售卖。 页面已增加防盗追踪，如有侵权极客邦将依法追究其法律责任。

上一篇 16 | 验证爬虫：我到底要不要百分百投入？

下一篇 18 | 如何搞定老板（上）：如何编造虚无缥缈的OKR？

精选留言 (1)

写留言

leslie

2022-03-09

先说第一个问题： 一个公司，发了一个 JD，要求应聘者同时做过爬虫和反爬虫，你认为他们实际上想招哪一个？ 1）这说明其实他不清楚真正想要什么，招一个进来探路的，二者技能兼备； 2）若单方面强，而实际又有所偏差；可以同样的招聘在招一个反向技能擅长的； 2.你和 HR 说要招一个爬虫工程师，HR 直接去网上搜了个 JD。这样最坏可能引发什么问题？ 最坏的可能就是引发竞争对手去招反爬工程师，而且根据你的需求去针对性...

展开

作者回复: 其实更可怕的是抄了个有pr风险的jd